
Journal of Urban Health: Bulletin of the New York Academy of Medicine Vol. 80, No. 2, Supplement 1 2003

Investigation of Disease Outbreaks Detected by "Syndromic" Surveillance Systems

Julie A. Pavlin

ABSTRACT *Syndromic surveillance systems can detect potential disease outbreaks quickly and can provide useful tools to assist in outbreak investigation. The steps used to investigate diseases detected through these newer methods are not that different from traditional investigative measures, but the differences and limitations of the systems must be understood. With syndromic surveillance systems, there is often readily available electronic demographic information that can help define the epidemic and direct disease control measures. The diagnosis needs to be confirmed as quickly as possible, however, as specific diagnostic information will be missing with early detection from nonspecific data. It is also important not to disregard smaller, nonsevere rises in disease incidence as they might be a harbinger of a worsening outbreak. The rapidity of most syndromic surveillance systems also requires an equally rapid response, and planning must be done to prioritize alert categories and the response sequence to best utilize the information available in these new systems.*

KEYWORDS *Epidemiology, Outbreak investigation, Syndromic surveillance.*

INTRODUCTION

Any unusual increase in disease incidence should be investigated. The intensity and effort of the investigation is dependent on the severity of the disease, the number of people affected, the potential for the disease to spread, and the effectiveness of available countermeasures.[1] Disease outbreaks detected through new health indicator surveillance systems, such as those termed *syndromic surveillance*, should be investigated similarly to those detected by traditional means. However, some steps in the investigative process differ slightly or have a different priority when an alert is generated by a syndromic surveillance system.

NEW INFECTIOUS DISEASE SURVEILLANCE TOOLS

With the growing awareness of the threat of emerging infections and bioterrorism, many new types of disease surveillance systems have been developed.[2–10] Some of the systems use medical data sources that are routinely collected for other purposes (e.g., emergency room logs), some collect new data at the point of patient encounter, and some use nonclinical data (e.g., pharmacy sales, school absenteeism) to trigger an alert. As these systems are implemented, there is growing concern of how to respond to alerts and how to investigate them based on information that has not traditionally been used by public health departments.

Dr. Pavlin is Chief, Department of Field Studies, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910. (E-mail: julie.pavlin@amedd.army.mil)

Syndromic surveillance systems often use data sources provided by clinicians, such as *International Classification of Diseases, 9th Revision* (*ICD-9*) diagnostic codes or basic symptom information. Unlike traditional systems, the information is provided on all patients for all encounters instead of concentrating on one disease or sample of the population. In many cases, the new surveillance tool might be more sensitive than traditional systems since all encounters are captured. However, with the increase in sensitivity comes a potential for false-positive alarms or the detection of minor, inconsequential increases in disease incidence. This can make the investigation of detected outbreaks from syndromic surveillance systems differ slightly from that of traditional alerts.

DISEASE OUTBREAK INVESTIGATION STEPS

There are essential steps in any epidemiological investigation regardless of how the outbreak is detected.[11] With the use of syndromic surveillance, some steps might receive greater emphasis than others, depending on various factors such as data sources. Using generally accepted steps in an outbreak investigation, the differences are highlighted.

Confirm Existence of Outbreak

First, establish the existence of an outbreak. This might be the most important, and possibly the most difficult, step with syndromic surveillance. By definition, these surveillance systems detect outbreaks early—possibly before patients present for medical care. Therefore, usual methods of outbreak confirmation, such as records reviews or patient interviews, might not be possible. There are other ways that an epidemiologist can look at available data to determine if further investigation is warranted. First, all available data sources should be correlated and checked for obvious causes of error or expected variation. Possible causes of false alerts include promotional sales at drug stores, use of incorrect diagnostic codes, or an increase in the baseline population under surveillance. If the outbreak is confirmed in multiple data sources, and there does not seem to be an alternate explanation for the alert, then a full investigation should follow.

Verify the Diagnosis

The second step is to verify the diagnosis. Verification might not occur immediately, and the investigation and any appropriate preventive measures should continue during this process. Syndromic surveillance can assist by alerting public health personnel to use laboratory tests such as viral cultures, rapid polymerase chain reaction, or enzyme-based tests at an earlier point in the outbreak sequence.

Estimate the Number of Cases

Once the diagnosis is verified to the best extent possible, a case definition should be developed to count the number affected. Traditionally, this case definition uses a combination of clinical and laboratory data if available. For example, a case definition for influenza might be the presence of a fever and sore throat or cough or a positive viral culture. With syndromic surveillance, the ability to create a case definition and count the number of people who meet that definition is easy—the difficulty lies in interpreting the counts.

For example, with data that use outpatient diagnostic codes, a case could be defined as any patient diagnosed with any acute respiratory illness code. The data-

base can be queried for these patients, and basic demographic information can be determined. However, what is not known is how many of these are true cases involved in the current epidemic, how many have similar, uninvolved illnesses, or how many are coding errors. Syndromic surveillance cannot eliminate the need for "shoe leather" epidemiology—tracking down the clinical information and exposure histories—but it can direct the investigation by rapidly providing electronic demographic information.

Orient the Data to Person, Place, and Time

Once the investigator has determined the potential number of cases, descriptive epidemiology can define the epidemic in terms of person, place, and time. With already collected demographic data, syndromic surveillance can facilitate definition of these terms much earlier in the epidemic curve and thus enable the institution of preventive measures based on these characteristics.

Develop and Evaluate Hypotheses

After the characteristics of the epidemic are defined, an epidemiologist can develop and test hypotheses on what is causing the outbreak and what the suspected risk factors are for acquiring the disease. These include the source of the etiologic agent and the mode of transmission. Once a hypothesis is generated, it can be evaluated by comparing it to known facts about the potential disease agent or through analytic epidemiological methods, such as cohort or case-control studies. These tests can be performed with data generated by syndromic surveillance systems.

Institute Control Measures and Communicate Findings

The final steps in an outbreak investigation are to implement control measures based on the most likely hypothesis and to communicate the findings to other health professionals. Simple control measures can be based on data generated by early warning surveillance systems, such as information that certain geographic areas have been affected, that the illness appears to be mostly in school-aged children, or that it appears to be communicable. As more laboratory diagnostic information becomes available, the control measures can become more specific.

EXAMPLES OF OUTBREAKS DETECTED BY THE ELECTRONIC SURVEILLANCE SYSTEM FOR THE EARLY NOTIFICATION OF COMMUNITY-BASED EPIDEMICS

The Department of Defense Global Emerging Infections System (DoD-GEIS) has developed the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) as a tool to monitor the health status of military health care beneficiaries worldwide.[10] Using outpatient *ICD-9* diagnostic codes, a set of diagnoses or symptoms potentially caused by an infectious disease has been created by grouping similar codes. For example, respiratory, gastrointestinal, and unspecified febrile illnesses are all syndrome groups in ESSENCE. Since 1999, ESSENCE has been operating in the National Capital Area, and since 2001, it has been used in all military clinics worldwide. During this time, many unexpected increases in numbers of potentially infectious cases have been recorded. Using the steps outlined previously, some of these increases have quickly been determined not to be true outbreaks, but to be caused by coding errors, a recent influx in population, or a new provider's coding idiosyncrasies.

Other detected disease outbreaks include many that were probably true outbreaks, but were small, declined rapidly, and did not warrant detailed investigation. The alert of a possible epidemic did allow notification to providers in that area that something unusual might be occurring and to take cultures or notify local public health personnel if more cases presented.

Some outbreaks detected by ESSENCE have been investigated thoroughly, including a significant outbreak that occurred at a Marine Corps base in January 2002. More than 135 Marines were involved in a gastrointestinal outbreak during a 1-week period. The epidemic curve from ESSENCE is displayed in the Figure. Laboratory results eventually confirmed the hypothesis of a Norwalklike virus as the cause of the illness in this outbreak. Using this outbreak and the steps outlined above, the differences noted between this investigation and what would be expected from a traditional notification system investigation are described in Table 1.

SPECIFIC ISSUES IN INVESTIGATION OF A DISEASE OUTBREAK DETECTED BY SYNDROMIC SURVEILLANCE

Passive surveillance often relies on an astute clinician for notification of an acute outbreak. When this occurs, a clinician might request the assistance of public health personnel. With syndromic surveillance, public health might be mobilized first, with the epidemiologist contacting clinicians who possibly have no idea that there is a potential problem. The requirement for the public health system to mobilize clinicians instead of the reverse might pose a significant challenge. With the use of new surveillance systems, an appropriate alert sequence should be defined for a range of levels of seriousness, and written response procedures should be more varied than in the past.

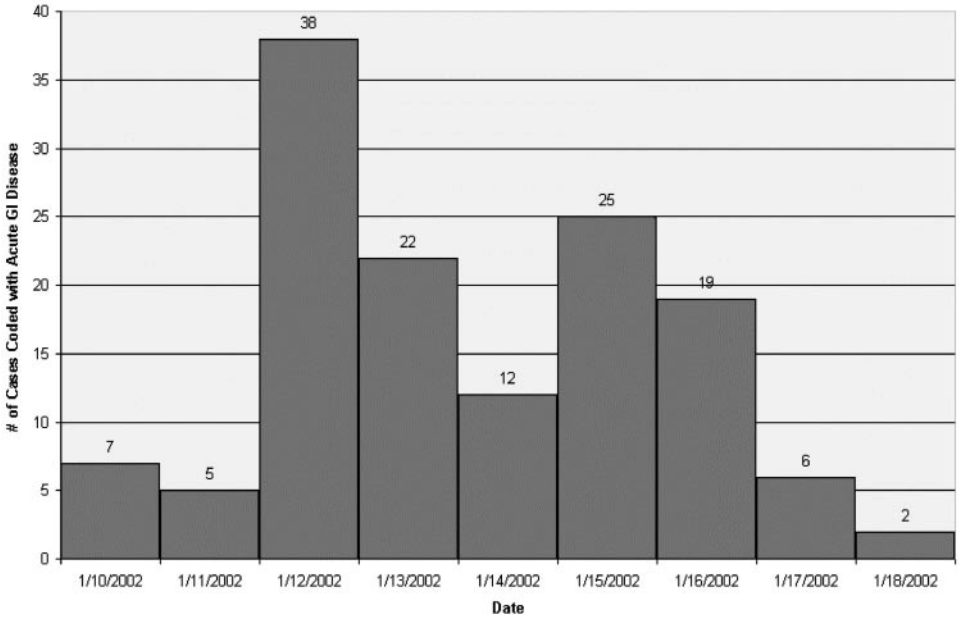

FIGURE. Example of gastrointestinal (GI) disease outbreak detected by ESSENCE.

TABLE 1. Differences in outbreak investigative steps using the January 2002 Marine Corps gastrointestinal outbreak as an example

Step	ESSENSE	Traditional
Confirm existence of outbreak	Used ESSENCE data to alert; number of cases much higher than baseline; confirmed outbreak through call to clinic.	Public health might have eventually been notified by the clinic; would still need to determine that cases were actually higher than normal.
Verify diagnosis	Stool samples (16) had been collected, but standard tests negative. Once ESSENCE detected the outbreak, the samples were preserved and sent for viral testing.	If brought to the attention of public health personnel, similar actions would have been taken.
Estimate the number of cases (case definition)	Used ESSENCE data with any *ICD-9* code for vomiting or diarrhea or combination. During records review, used the definition of onset of acute vomiting and/or diarrhea during the epidemic period. The ESSENCE database allowed immediate identification of those affected.	Would have used similar case definition for records review, although would not have had the ability to know immediately which records should be investigated first.
Orient to person, place and time	Easily generated epidemic curve and described with time, place, and person. Knew age groups affected and where individuals lived and when they were medically evaluated.	Could only do this after extensive and time-consuming records review.
Develop and evaluate hypotheses	Based on explosive nature and potential person-to-person spread, hypothesis of Norwalklike virus as causative agent. Fit all epidemiological variables.	After notification and records review, would have hypothesized similar cause.
Implement control measures	ESSENCE data led to belief of person-to-person spread and potential source (ill food handler) found on records review. Therefore, reinforced training for food handlers and recommended separation of ill Marines from crowded living conditions.	Same recommendations, probably delayed due to time required to determine outbreak characteristics.
Communicate findings	Communicated findings to public health personnel on base and presented at international scientific meetings.	Same.

Another issue is what to do with small outbreaks of questionable public health importance. Syndromic systems will often detect rises in various syndromic categories that might be validated by similar findings in multiple data sources. These outbreaks might not involve many patients or be caused by severe illnesses. These are not false positives, but beyond initial analysis of available demographic data, they might not be worth using limited resources to investigate. For these reasons, predefined prioritized response actions should be developed.

The detection of early changes in disease incidence can also be very problematic. One reason for using syndromic surveillance systems is for early detection of naturally occurring or bioterrorist-caused disease epidemics. If detected early, the premise is that something can be done to mitigate the situation. Since most potential bioterrorist weapons produce very nonspecific symptoms early in the course of disease, a system that detects an increase in nonspecific infectious disease symptoms could be an early warning important to capture. However, if it is decided that the increase is too small or too mild to investigate, then the goal of detecting outbreaks early has been defeated. Each regional system should develop guidelines to assist the epidemiologist in deciding whether to pursue an investigation. The guidelines will differ depending on the demographics of the region and the data sources available. Some potential recommendations are listed in Table 2.

Investigations triggered by syndromic surveillance systems can provide the power to conduct multiple steps in the investigation simultaneously, rapidly, and efficiently. For example, there might be the capacity to look at baseline historic data and patterns, to apply predefined models to geographic patterns, and to contact patients for follow-up interviews. Investigations will probably be facilitated by a nontraditionally heavy emphasis on predefined information technology tools. The

TABLE 2. Potential epidemiological factors that call for increased investigation or monitoring

Factors	Reason
Disease located in one geographic area	Might indicate a point source of a disease agent that can be discovered and controlled
Severe symptoms/diagnoses such as encephalitis or death	Indicates disease process that needs rapid investigation due to severity
Rapid rise to very high numbers of illness two to three times normal baseline with steep epidemic curve	Potential for continuing rapid rise in numbers; requires immediate investigation to institute control measures
Outbreak detected and confirmed in multiple data sources	Unlikely to be due to error; possibly widespread
Outbreak occurring at an unusual time or place (e.g., respiratory/influenzalike symptoms in the summer)	Might indicate potential bioterrorist attack or unexpected disease introduction
Outbreak confined to one age or gender group	Might indicate targeted population or early signs in a susceptible population (very young or very old)
Number of cases continuing to rise over time	Indicates sustained outbreak that might continue to grow

relationship between surveillance and outbreak investigation should be more seamless than traditionally has been the case.

Final important issues when investigating a potential disease outbreak detected with alternate data sources are privacy and legal concerns. Most systems use data that were not created for surveillance purposes. Some systems link to personal identifiers that can be of great use when performing an outbreak investigation, but require stringent data-handling procedures. It is very important to understand the legal limitations and requirements in safeguarding this information and using it only for permitted public health reasons. Even data that are not part of a personal medical record, such as pharmacy sales, can be subject to legal stipulations based on the commercial privacy needs of the data source.

ACKNOWLEDGEMENT

The Department of Defense Global Emerging Infections System and the Defense Applied Research Projects Agency funded investigations and evaluations performed for the ESSENCE system.

The views expressed are those of the author and do not reflect the position of the US Army or the Department of Defense.

Kendall Brown, Virginia Foster, Jay Mansfield, Eugene Elbert, Jeff Gunzenhauser, Ken Cox, and Patrick Kelley were of assistance in evaluating many of the outbreaks detected by ESSENCE. Joe Lombardo, Howard Burkom, Rich Wojcik, Sheri Lewis, Steve Magruder, Wayne Loschen, and Scott Mauer from the Johns Hopkins University Applied Physics Laboratory provided technical support and expertise that allowed the visualization and detection of outbreaks in the ESSENCE system.

REFERENCES

1. Reingold AL. Outbreak investigations—a perspective. *Emerg Infect Dis*. 1998;4:21–27.
2. Miller JR, Mikol Y. Surveillance for diarrheal disease in New York City. *J Urban Health*. 1999;76:388–390.
3. Pavlin JA, Kelley PW, Mostashari F, et al. Innovative surveillance methods for monitoring dangerous pathogens. In: Institute of Medicine (US). *Biological Threats and Terrorism: Assessing the Science and Response Capabilities*. Washington, DC: National Academy of Sciences; 2002:185–196.
4. Espino JU, Tsui F-C, Wagner M. Realtime outbreak detection system (RODS). Available at: www.health.pitt.edu/rods/rods.htm. Accessed February 11, 2002.
5. Lazarus R, Kleinman KP, Dashevsky I, DeMaria A, Platt R. Using automated medical records for rapid identification of illness syndromes (syndromic surveillance); the example of lower respiratory infection. *BMC Public Health*. 2001;1:9.
6. New Mexico Department of Health. Rapid Syndrome Validation Project (RSVP) project description. Available at: http://epi.health.state.nm.us/rsvpdesc/default.asp. Accessed February 11, 2002.
7. Lober WB, Karras BT, Wagner MM, et al. Roundtable on bioterrorism detection: information system-based surveillance. *J Am Med Inform Assoc*. 2002;9:105–115.
8. Green MS, Kaufman Z. Surveillance for early detection and monitoring of infectious disease outbreaks associated with bioterrorism. *Isr Med Assoc J*. 2002;4:503–506.
9. Centers for Disease Control and Prevention. Syndromic surveillance for bioterrorism following the attacks on the World Trade Center—New York City, 2001. *MMWR Morb Mortal Wkly Rep*. 2002;51:13–15.

10. Lewis MD, Pavlin JA, Mansfield JL, et al. Disease outbreak detection system using syndromic data in the greater Washington, DC area. *Am J Prev Med*. 2002;23:180–186.
11. Tyler CW Jr, Last JM. Epidemiology. In: Wallace RB, ed. *Maxcy-Rosenau-Last Public Health and Preventive Medicine*. 14th ed. East Norwalk, CT: Appleton and Lange; 1998:5–33.